PARAMETER ESTIMATION OF THE BOUNDED BINOMIAL DISTRIBUTION

by

HO Yat Fan

A

Thesis

Submitted to

the Graduate School of

The Chinese University of Hong Kong

(Division of Statistics)

In Partial Fulfillment

of the Requirement for the Degree of

Master of Philosophy (M. Phil.)

May, 1983

THE CHINESE UNIVERSITY OF HONG KONG

GRADUATE SCHOOL

The undersigned certify that we have read a thesis, entitled " PARAMETER ESTIMATION OF THE BOUNDED BINOMIAL DISTRIBUTION " submitted to the Graduate School by HO Yat Fan (何一帆) in partial fulfillment of the requirement for the degree of Master of Philosophy in Statistics. We recommend that it be accepted.

Lam HK

Dr. H.K. Lam, Supervisor

N.N. Chan

Dr. N.N. Chan

Lee Sik Yum

Dr. S.Y. Lee

Professor L.R. Shenton
External Examiner

Acknowledgements

I wish to express my sincere thanks to my advisor Dr. H.K. Lam, whose encouragement, stimulation, patience, valuable advice and excellent guidance have made this thesis possible. I am extremely grateful to Professor H. Tong, Dr. N.N. Chan, Dr. S.Y. Lee and Dr. C.Y. Leung for their substantial assistance during my two years of graduate studies. I would also like to thank Dr. N.N. Chan and Dr. S.Y. Lee for reading this thesis and give their helpful suggestions. Finally I would like to thank my friend, Mr. W.K. Fong, for his proof-reading of this manuscript.

Abstract

The purpose of this thesis is to give a discussion of a modified binomial distribution, the Bounded Binomial Distribution. In our daily life, there are many situations in which the number of occurrences x of the events are bounded by a number t which is less than the total number of trials. Hence the usual binomial distribution, with parameters n and p say, may not be applicable. To give a clear picture, we give some concrete examples in Chapter 1. It is obvious that we can subdivide our discussion into three cases, namely the Bounded Above Binomial Distribution, the Bounded Below Binomial Distribution and the Doubly Bounded Binomial Distribution. However, we confine our investigation to the Bounded Above Binomial Distribution only. The other types of Bounded Binomial Distribution can be defined similarly. In Chapter 2, we shall discuss the distribution of the Bounded Above Binomial Distribution, giving its p.d.f., moments and generating functions. Chapter 3 takes care of the existence of sufficient statistic and minimum variance bound estimator. In Chapters 4 and 5, we give a close investigation of the maximum likelihood estimator, moment estimator and some other estimators of the parameter p. To conclude our discussion, we outline a brief account of comparison of the various estimates of p by using computer simulation (with fixed n and the bound t). Many interesting numerical examples are given in this chapter in supporting our discussion. In addition, many extreme cases are also taken into consideration.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION

Experimenters and decision-makers are frequently faced with dichotomous situations. For example, a survey undertaken before the marketing of a new product elicits either a favorable response from each individual sampled; inspection of the output of a production run results in either defective or nondefective items; and so on. The earliest studies of the probability aspects of a dichotomy were undertaken by James Bernoulli in the 17th century.

Consider an experiment $\mathcal{E}$ and let A be some event associated with $\mathcal{E}$. Suppose that $P(A) = p$ and $P(\overline{A}) = 1-p$. Consider n independent repetitions of $\mathcal{E}$. Assume that $P(A) = p$ remains the same for all repetitions. Let the random variable X be defined as follow: X = number of times the event A occurred. We call X a binomial random variable with parameters n and p. Its probability density function is

$$\Pr(X=k) = \binom{n}{k} p^k (1-p)^{n-k} , \qquad k = 0, 1, 2, \ldots , n . \tag{1.1}$$

The binomial distribution has been widely used since it was introduced. It serves as an excellent probability model in a number of experimental situations. Examples of applications are numerous (see Hogg and Tanis (1977) and Hymans, S.H. (1967)).

However, there are many situations in which the number of occurrences of the events are bounded by a number, say t, which is less than the total number of trials, thus the binomial distribution may not be applicable.

For example, in a shooting training course, each person is required to execute n shots towards a fixed target. It is obvious that the total number of shots strike the target in $x = 0, 1, 2, \ldots, n$ number of times. Suppose the target will ruin if it suffers t shots or more. Hence it is impossible to distinguish whether there are t shots, t+1 shots, ... , or n shots embedded in it. As a result, we shall treat $x = t, t+1, \ldots, n$ as the same group. Consider another example, suppose there are n persons who may book the front stall tickets of a particular theatre. The number of seats in the front stall of this theatre is t with $t < n$. It is clear that if the t seats are filled, it is impossible to distinguish whether there are t, t+1, ... , or n persons attempt to book the seat in the front stall. Again, we treat those indistinguishable ones as one group.

Lam (1982) investigated the multinomial case. Consider a sequence of n trials each of which results in one and only one of s mutually exclusive events $A_1, A_2, \ldots, A_s$ with probabilities $p_1, p_2, \ldots, p_s$ respectively. If X_i denotes the number of occurrences of A_i in the sequence of n trials, $i = 1, 2, \ldots, s$, then the random vector $(X_1, \ldots, X_s)$ has the usual multinomial distribution (see Johnson and Kotz (1969)) with parameters $n, p_1, \ldots, p_s$. Its p.d.f. is given by

$$\Pr\left[\bigcap_{i=1}^{s} (X_i=x_i)\right] = n! \prod_{i=1}^{s} (p_i^{x_i}/x_i!) , \quad 0 \le x_i \le n , \quad \sum_{i=1}^{s} x_i = n . \qquad (1.2)$$

Lam considered the following modification. Let t_i denotes the bound of X_i, thus the frequency X_i cannot exceed t_i, $i = 1, 2, \ldots, s$. Of course

$t_i < n$ for at least one i. With such a restriction, the sequence of the trials is no longer independent. If at any stage during the process of the trials, the number of occurrences of a particular event, A_j say, already attains the corresponding bound, then A_j will be excluded from the set of possible events in the remaining trials. Thus in the following trials, only those events with the corresponding bounds still unattained may occur. It would be plausible to assume that the probabilities of occurrence of these bound-unattained events remain in the same ratios as their initial probabilities. Then each will be enhanced by a factor $1/(1 - \sum p_i)$, in which the summation sign sums over those events having the bound already been attained. Therefore, for the given initial probabilities $p_1, p_2, \ldots, p_s$ of the events $A_1, A_2, \ldots, A_s$ respectively, the probability of occurrence of A_i on a particular trial is

(i) zero, if the number of occurrence of A_i already attains t_i;

or (ii) $\dfrac{p_i}{(1 - \sum p_j)}$, if at that moment it happens that the current values of some X_j's already attain the corresponding bounds, and $\sum p_j$ is the sum of the corresponding initial probabilities;

or (iii) p_i, if at that moment the current values of all the X_j's are less than the corresponding t_j's.

Under such conditions, the joint distribution of $(X_1, \ldots, X_s)$ at the end of the sequence of n trials will be called a Bounded Multinomial Distribution. Lam (1982) gave the joint probability density function of $X_1, \ldots, X_s$.

It is clear the usual multinomial distribution is the special case that $t_i \geq n$, for all i's. Also notice that if $x_i < t_i$, for all i's, then the joint probability of $X_i = x_i$, $i = 1, 2, \ldots, s$ is (1.2), the same as that given by the usual multinomial distribution.

We are interested in the binomial case, i.e. $s = 2$. It is the distribution of X_1, the number of occurrence of a particular event, A_1 say. For simplicity we will write X_1 as X. Reduced from Lam (1982)'s (2.3), the p.d.f. of X is

$$\Pr(X = x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & \text{if } 0 \leq x < t, \\ \sum_{j=t}^{n} \binom{j-1}{t-1} p^t (1-p)^{j-t}, & \text{if } x = t. \end{cases} \qquad (1.3)$$

Since $X \leq t$, we will call X a bounded above binomial variable with parameters n, p and t.

In Chapter 2, we shall discuss the distribution. Chapter 3 contains a general discussion of existence of sufficient statistic and minimum variance bound estimator. Chapters 4 and 5 deal with maximum likelihood estimate, moment estimate and other estimates of the parameter p with n and t fixed. In Chapter 6, comparison of the various estimates of the parameter p by using computer simulation is discussed. The last Chapter is summaries and conclusions.

CHAPTER 2. THE DISTRIBUTION

2.1 The p.d.f. of Bounded Binomial Distribution

First of all, let us find two alternative expressions for $\sum_{j=t}^{n} \binom{j-1}{t-1} p^t (1-p)^{j-t}$ which appears in (1.3). It can be shown that

$$\sum_{j=t}^{n} \binom{j-1}{t-1} p^t (1-p)^{j-t} = \sum_{j=t}^{n} \binom{n}{j} p^j (1-p)^{n-j} , \qquad (2.1)$$

$$= \int_0^p \frac{\Gamma(n+1)}{\Gamma(t)\Gamma(n-t+1)} (1-x)^{n-t} x^{t-1} dx . \qquad (2.2)$$

where $\Gamma(\cdot)$ is the gamma function.

We prove (2.1) by mathematical induction.

For $t = 1$ $\quad$ L.H.S. $= \sum_{j=1}^{n} \binom{j-1}{0} p^1 (1-p)^{j-1}$

$$= p \sum_{j=1}^{n} (1-p)^{j-1}$$

$$= p \frac{1-(1-p)^n}{1-(1-p)}$$

$$= 1 - (1-p)^n .$$

R.H.S. $= \sum_{x=1}^{n} \binom{n}{x} p^x (1-p)^{n-x}$

$$= 1 - \sum_{x=0}^{0} \binom{n}{x} p^x (1-p)^{n-x}$$

$$= 1 - (1-p)^n .$$

Thus $\quad$ L.H.S. = R.H.S.

Suppose (2.1) is true for t, that is

$$\sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x} = \sum_{j=t}^{n} \binom{j-1}{t-1} p^t(1-p)^{j-t} \quad . \tag{2.3}$$

To show that (2.1) is true for t + 1 as long as $n \geq t+1$, we notice that

$$\binom{j}{t} = \binom{j-1}{t-1} + \binom{j-1}{t} \quad . \tag{2.4}$$

Hence

$$\text{L.H.S.} = \sum_{j=t+1}^{n} \binom{j-1}{t+1-1} p^{t+1}(1-p)^{j-(t+1)}$$

$$= \sum_{j=t}^{n} \binom{j-1}{t} p^{t+1}(1-p)^{j-(t+1)} \qquad \text{since} \binom{t-1}{t} = 0$$

$$= \sum_{j=t}^{n} \binom{j}{t} p^{t+1}(1-p)^{j-(t+1)} - \sum_{j=t}^{n} \binom{j-1}{t-1} p^{t+1}(1-p)^{j-(t+1)} \quad \text{(by (2.4))}$$

$$= \sum_{i=t+1}^{n+1} \binom{i-1}{t} p^{t+1}(1-p)^{i-1-(t+1)} - \frac{p}{1-p} \sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x} \quad \text{(by (2.3))}$$

$$= \frac{1}{1-p} \sum_{i=t+1}^{n+1} \binom{i-1}{t} p^{t+1}(1-p)^{i-(t+1)} - \frac{p}{1-p} \sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x}$$

$$= \frac{1}{1-p} \left[\sum_{i=t+1}^{n} \binom{i-1}{t} p^{t+1}(1-p)^{i-(t+1)} + \binom{n}{t} p^{t+1}(1-p)^{n-t} \right]$$

$$- \frac{p}{1-p} \sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x}$$

$$= \frac{1}{1-p}(\text{L.H.S.}) + \frac{1}{1-p}\binom{n}{x} p^{t+1}(1-p)^{n-t} - \frac{p}{1-p} \sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x}$$

or $$\text{L.H.S.} \left[\frac{1}{1-p} - 1 \right] = \frac{p}{1-p} \sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x} - \frac{1}{1-p}\binom{n}{x} p^{t+1}(1-p)^{n-t}$$

or $\quad \text{L.H.S.} = \sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x} - \frac{1}{p}\binom{n}{t} p^{t+1}(1-p)^{n-t}$

$$= \sum_{x=t+1}^{n} \binom{n}{x} p^x(1-p)^{n-x} + \binom{n}{t} p^t(1-p)^{n-t} - \binom{n}{t} p^t(1-p)^{n-t}$$

$$= \text{R.H.S.}$$

Thus (2.1) is established.

To see that (2.2) is valid, let $x = p - pz$

$$\int_0^p \frac{\Gamma(n+1)}{\Gamma(t)\Gamma(n-t+1)}(1-x)^{n-t}x^{t-1}dx = \frac{\Gamma(n+1)}{\Gamma(t)\Gamma(n-t+1)}\int_1^0 \left[(1-p)+pz\right]^{n-t}p^t(1-z)^{t-1}(-1)dz$$

$$= \frac{\Gamma(n+1)}{\Gamma(t)\Gamma(n-t+1)}\, p^t \int_0^1 \sum_{i=0}^{n-t}\binom{n-t}{i}(pz)^i(1-p)^{n-t-i}(1-z)^{t-1}dz$$

$$= \frac{\Gamma(n+1)}{\Gamma(t)\Gamma(n-t+1)}\, p^t \sum_{i=0}^{n-t}\binom{n-t}{i} p^i(1-p)^{n-t-i}\int_0^1 z^i(1-z)^{t-1}dz$$

$$= \sum_{i=0}^{n-t} \frac{n!}{(t-1)!(n-t)!}\,\frac{(n-t)!}{i!(n-t-i)!}\, p^{t+i}(1-p)^{n-t-i}\,\frac{(t+1)!i!}{(t+i)!}$$

$$= \sum_{i=0}^{n-t}\binom{n}{t+i} p^{t+i}(1-p)^{n-t-i}$$

$$= \sum_{j=t}^{n}\binom{n}{j} p^j(1-p)^{n-j} \quad .$$

Since the quantity on the right hand side of (2.1) is the upper tail probability of a binomial distribution, we shall denote the quantities in (2.1) and (2.2) as $B(n,p;t^+)$. Thus the p.d.f. of a bounded above binomial variate with parameters n, p and t can be written as

$$f(x;p) = \begin{cases} \binom{n}{x} p^x(1-p)^{n-x} , & \text{if } 0 \leq x < t , \\ B(n,p;t^+) , & \text{if } x = t . \end{cases} \tag{2.5}$$

(2.5) reveals the characteristic of the bounded above binomial distribution that it degenerates the upper tail probabilities of a binomial distribution onto a single point at t.

We should distinguish it from the truncated binomial distribution. The truncated binomial distribution is a kind of conditional distribution. A singly truncated binomial distribution is formed if only the values 0, 1,..., (r_1-1) $(r_1 \geq 1)$ or the values $N-r_2+1, \ldots , N$ $(r_2 \geq 1)$ are omitted. The distribution formed by omission of the value 0 only, giving

$$\Pr(x = k) = \frac{\binom{N}{k} p^k q^{N-k}}{1 - q^N} \qquad k = 1, 2, \ldots , N \tag{2.6}$$

is sometimes called the positive binomial distribution. (see Johnson and Kotz (1969))

(2.5) suggests the definition of other types of bounded binomial distribution, namely, the Bounded Below Binomial Distribution and the Doubly Bounded Binomial Distribution. They can also serve as probability models for many real life situations. In like manner, the p.d.f. of a Bounded Below Binomial Distribution is

$$f(x) = \begin{cases} B(n,p;t^-) , & \text{if } x = t , \\ \binom{n}{x} p^x(1-p)^{n-x} , & \text{if } t < x \leq n , \end{cases}$$

where $\quad B(n,p,;t^-) = \sum_{j=o}^{t} \binom{n}{j} p^j(1-p)^{n-j}$;

and the p.d.f. of a Doubly Bounded Binomial Distribution is

$$f(x) = \begin{cases} B(n,p;t_2^-) , & \text{if } x = t_2 , \\ \binom{n}{x} p^x(1-p)^{n-x} , & \text{if } t_2 < x < t_1 , \\ B(n,p;t_1^+) , & \text{if } x = t_1 . \end{cases}$$

In fact, it is clear that the Bounded Above and the Bounded Below Binomial Distribution are mirror image of each other. Therefore in the sequel we shall consider the Bounded Above Binomial Distribution only.

2.2 Moments and Generating Functions

2.2.1 Probability generating function

The probability generating function of a Bounded Above Binomial Distribution with parameters n, p and t is

$$\begin{aligned} g(\theta) &= E(\theta^x) \\ &= \sum_{x=0}^{t-1} \binom{n}{x} p^x(1-p)^{n-x}\theta^x + \theta^t \sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} (pe)^x(1-p)^{n-x} + \sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x}\left[\theta^t - \theta^x\right] \\ &= \left[(1-p) + p\theta\right]^n + \sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x}\left[\theta^t - \theta^x\right] . \qquad (2.7) \end{aligned}$$

2.2.2 Characteristic function

The characteristic function of this distribution is

$$\rho(\theta) = E(e^{i\theta x})$$

$$= \sum_{x=0}^{t-1} e^{i\theta x} \binom{n}{x} p^x(1-p)^{n-x} + e^{i\theta t} \sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x}$$

$$= \sum_{x=0}^{t-1} \binom{n}{x} (e^{i\theta}p)^x(1-p)^{n-x} + \sum_{x=t}^{n} \binom{n}{x} e^{i\theta t}p^x(1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (e^{i\theta}p)^x(1-p)^{n-x} - \sum_{x=t}^{n} \binom{n}{x} (e^{i\theta}p)^x(1-p)^{n-x}$$

$$+ \sum_{x=t}^{n} \binom{n}{x} e^{i\theta t}p^x(1-p)^{n-x}$$

$$= \left[(1-p) + pe^{i\theta}\right]^n + \sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x}\left[e^{i\theta t} - e^{i\theta x}\right] \quad . \qquad (2.8)$$

2.2.3 Moment generating function

The moment generating function can easily be established. Replace $i\theta$ in (2.8) by θ, we obtain

$$M(\theta) = E(e^{\theta x})$$

$$= \left[(1-p) + pe^{\theta}\right]^n + \sum_{x=t}^{n} \binom{n}{x} p^x(1-p)^{n-x}\left[e^{\theta t} - e^{\theta x}\right] . \qquad (2.9)$$

Also we notice that $g(e^{\theta}) = M(\theta)$.

2.2.4 Moments

(a) rth factorial moment:

Let $x^{(r)} = x(x-1) \ldots (x-r+1)$, then expressions for moments about zero in terms of descending factorial moments is

$$\mu_{(r)}(x) = E\left[x^{(r)}\right]$$

$$= \sum_{x=0}^{t-1} x^{(r)} \binom{n}{x} p^x(1-p)^{n-x} + t^{(r)}B(n,p;t^+)$$

$$= \sum_{x=0}^{t-1} x(x-1)\ \ldots\ (x-r+1) \frac{n!}{x!(n-x)!} p^x(1-p)^{n-x} + t^{(r)}B(n,p;t^+)$$

$$= \sum_{x=r}^{t-1} x(x-1)\ \ldots\ (x-r+1) \frac{n!}{x(x-1)\ \ldots\ (x-r+1)(x-r)!(n-x)!} p^x(1-p)^{n-x}$$

$$+ t^{(r)}B(n,p;t^+)$$

$$= n(n-1)\ \ldots\ (n-r+1)\, p^r \sum_{x=r}^{t-1} \binom{n-r}{x-r} p^{x-r}(1-p)^{n-x} + t^{(r)}B(n,p;t^+)$$

$$= n^{(r)}p^r \sum_{y=0}^{t-1-r} \binom{n-r}{y} p^y(1-p)^{(n-r)-y} + t^{(r)}B(n,p;t^+)$$

$$= n^{(r)}p^r\, B\left[n-r,p;(t-1-r)^-\right] + t^{(r)}B(n,p;t^+)\ .$$

Hence

$$\mu_{(r)}(x) = \begin{cases} n^{(r)}p^r\, B\left[n-r,p;(t-1-r)^-\right] + t^{(r)}B(n,p;t^+)\ , & \text{if } r < t\ , \\ t^{(r)}B(n,p;t^+)\ , & \text{if } r = t\ , \\ 0\ , & \text{if } r > t\ . \end{cases} \tag{2.10}$$

(b) rth non-central moment about zero:

Denotes the rth non-central moment about zero by $\mu'_r(x)$, then

$$\mu'_r(x) = E(x^r)\ .$$

Since
$$x^r = \sum_{j=0}^{r} x^{(j)} \mathcal{S}_r^{(j)}$$

where $\mathcal{S}_r^{(j)} = \left.\frac{\Delta^j x^r}{j!}\right|_{x=0}$ is called a Stirling number of the second kind and Δ is the difference operator defined as $\Delta f(x) = f(x+1) - f(x)$. Hence

$$\mu'_r(x) = \sum_{j=0}^{r} \mu_{(j)}(x) \mathcal{S}_r^{(j)}\ . \tag{2.11}$$

In particular

$$\mu_1'(x) = \mu_{(1)}(x) = np\ B[n-1,p;(t-2)^-] + t\ B(n,p;t^+)$$

$$\mu_2'(x) = \mu_{(1)}(x)\mathcal{S}_2^{(1)} + \mu_2(x)\mathcal{S}_2^{(2)}$$

$$= np\ B[n-1,p;(t-2)^-] + t\ B(n,p;t^+)$$

$$+ n(n-1)p^2\ B[n-2,p;(t-3)^-] + t(t-1)\ B(n,p;t^+)$$

$$= np\ B[n-1,p;(t-2)^-] + n(n-1)p^2\ B[n-2,p;(t-3)^-] + t^2\ B(n,p;t^+)$$

$$\mu_3'(x) = \mu_{(1)}(x)\mathcal{S}_3^{(1)} + \mu_{(2)}(x)\mathcal{S}_3^{(2)} + \mu_{(3)}(x)\mathcal{S}_3^{(3)}$$

$$= np\ B[n-1,p;(t-2)^-] + 3n(n-1)p^2\ B[n-2,p;(t-3)^-]$$

$$+ n(n-1)(n-2)p^3\ B[n-3,p;(t-4)^-]$$

$$+ [t + 3t(t-1) + t(t-1)(t-2)]\ B(n,p;t^+)$$

$$= np\ B[n-1,p;(t-2)^-] + 3n(n-1)p^2\ B[n-2,p;(t-3)^-]$$

$$+ n(n-1)(n-2)p^3\ B[n-3,p;(t-4)^-] + t^3\ B(n,p;t^+)$$

$$\mu_4'(x) = \mu_{(1)}(x)\mathcal{S}_4^{(1)} + \mu_{(2)}(x)\mathcal{S}_4^{(2)} + \mu_{(3)}(x)\mathcal{S}_4^{(3)} + \mu_{(4)}(x)\mathcal{S}_4^{(4)}$$

$$= np\ B[n-1,p;(t-2)^-] + 7n(n-1)p^2\ B[n-2,p;(t-3)^-]$$

$$+ 6n(n-1)(n-2)p^3\ B[n-3,p;(t-4)^-]$$

$$+ n(n-1)(n-2)(n-3)p^4\ B[n-4,p;(t-5)^-] + t^4\ B(n,p;t^+)$$

and

$$\mathrm{var}(x) = \mu_2'(x) - \mu_1'^2(x)$$

$$= np\ B[n-1,p;(t-2)^-] + n(n-1)p^2\ B[n-2,p;(t-3)^-]$$

$$+ t^2\ B(n,p;t^+) - n^2p^2\ B^2[n-1,p;(t-2)^-]$$

$$-2npt\ B[n-1,p;(t-2)^-]\ B(n,p;t^+) - t^2\ B^2(n,p;t^+)\ .$$

CHAPTER 3. EXISTENCE OF SUFFICIENT STATISTIC AND MINIMUM VARIANCE BOUND ESTIMATOR

We are interested to estimate p assuming fixed value of n and t. Before going into deep detail discussion, we first consider the existence of sufficient statistic and minimum variance bound estimator.

Consider a family $\{f(x;\theta) : \theta \in \Omega\}$ of probability density functions where Ω is the interval set $\Omega = \{\theta, \gamma < \theta < \delta\}$, where γ and δ are known constants, and where

$$\begin{aligned} f(x;\theta) &= \exp\left[P(\theta)K(x) + S(x) + Q(\theta)\right] , \quad && a < x < b ; \\ &= 0 , && \text{elsewhere} . \end{aligned} \tag{3.1}$$

A probability density function of the form (3.1) is said to be a member of the exponential class of probability density function of the continuous type. If in addition

(a) neither a nor b depends upon θ, $\gamma < \theta < \delta$,

(b) $P(\theta)$ is a nontrivial continuous function of θ, $\gamma < \theta < \delta$,

(c) each of $K'(x) \not\equiv 0$ and $S(x)$ is a continuous function of x, $a < x < b$,

we say that we have a regular case of the exponential class.

Likewise, a p.d.f.

$$\begin{aligned} f(x;\theta) &= \exp\left[p(\theta)K(x) + S(x) + q(\theta)\right] , \quad && x = a_1, a_2, \ldots , \\ &= 0 , && \text{elsewhere} \end{aligned}$$

is said to represent a regular case of the exponential class of probability density functions of discrete type if

(a) the set $\{x;\ x = a_1, a_2, \dots\}$ does not depend upon θ,

(b) $n(\theta)$ is a nontrivial continuous function of θ, $\gamma < \theta < \delta$,

(c) $K(x)$ is a nontrivial function of x on the set $\{x;\ x = a_1, a_2, \dots\}$.

Let L be the likelihood function and $\tau(\theta)$ is some function of θ with an unbiased estimator u. From Kendall and Stuart (Chapter 17), under regularity, the necessary and sufficient condition for the existence of minimum variance bound (MVB) estimator is

$$\frac{\partial \log L}{\partial \theta} = A(\theta)\left\{u - \tau(\theta)\right\} , \tag{3.2}$$

where A is independent of the observations but may be a function of θ and u is a MVB estimator of $\tau(\theta)$.

Continuing to write $A(\theta)$ for the integral of the arbitrary function $A(\theta)$ in (3.2), the necessary form for the likelihood function is

$$\log L = uA(\theta) + P(\theta) + R(x_1, x_2, \dots , x_n) , \tag{3.3}$$

which we may re-write in the frequency-function form

$$f(x;\theta) = \exp\left\{A(\theta)B(x) + C(x) + D(\theta)\right\} , \tag{3.4}$$

where $u = \sum_{i=1}^{n} B(x_i)$, $R(x_1, x_2, \dots , x_n) = \sum_{i=1}^{n} C(x_i)$ and $P(\theta) = nD(\theta)$.

Thus, under regularity conditions, a distribution possesses MVB estimator if and only if the distribution is of exponential class. Now, write (2.5) as

$$f(x;p) = \begin{cases} \binom{n}{x} p^x(1-p)^{n-x} = \exp\left[(\log\frac{p}{1-p})x + \log\binom{n}{x} + n\log(1-p)\right], & 0 \le x < t , \\ B(n,p;t^+) = \exp\left[0 + 0 + \log B(n,p;t^+)\right], & x = t . \end{cases} \tag{3.5}$$

Clearly (3.5) is not of exponential class. However, we notice that

(a) the set $\{x: x = 0, 1, 2, \ldots, t\}$ does not depend upon p ,

(b) $F(p) = \log\left(\frac{p}{1-p}\right)$ is a nontrivial continuous function of p where $0 < p < 1$,

(c)
$$K(x) = \begin{cases} x , & \text{if } 0 \leq x < t ; \\ 0 , & \text{if } x = t , \end{cases}$$

is a nontrivial function of x on the set $\{x: x = 0, 1, \ldots, t\}$,

(d)
$$S(x) = \begin{cases} \log \binom{n}{x} , & \text{if } 0 \leq x < t ; \\ 0 , & \text{if } x = t , \end{cases}$$

(e)
$$Q(p,x) = \begin{cases} n \log (1-p) , & \text{if } 0 \leq x < t ; \\ \log B(n,p;t^+) , & \text{if } x = t , \end{cases}$$

which ensures that the Bounded Above Binomial Distribution satisfies the regularity conditions. Since $f(x;p)$ is not a member of exponential class, it follows that there is no MVB estimator for any function p.

Again from Kendall and Stuart (Chapter 17), we see that if u is sufficient for θ in a sample of n independent observations

$$\frac{\partial \log L}{\partial \theta} = \sum_{j=1}^{n} \frac{\partial \log f(x_j|\theta)}{\partial \theta} = K(u,\theta) , \tag{3.6}$$

where K is some function of u and θ. This equation remains true when we regard it as an equation in u for any fixed value of θ. Put

$$u = M\left\{\sum_{j=1}^{n} k(x_j)\right\} = M(\omega) , \tag{3.7}$$

where $\omega = \sum_{j=1}^{n} k(x_j)$ and M and k are arbitrary functions. Thus $K(M(\omega),\theta)$ is a function of ω and θ only, say $N(\omega,\theta)$. We have then, from (3.6), if the derivatives exist,

$$\frac{\partial^2 \log L}{\partial\theta\,\partial x_j} = \frac{\partial N}{\partial\omega}\frac{\partial\omega}{\partial x_j} \quad . \tag{3.8}$$

Now the left-hand side of (3.8) is a function of θ and x_j only and $\frac{\partial\omega}{\partial x_j}$ is a function of x_j only. Hence $\frac{\partial N}{\partial\omega}$ is a function of θ and x_j only. But it must be symmetrical in the x's, since ω is, and hence is a function of θ only. Hence integrating (3.8) with respect to ω, we get

$$N(\theta,\omega) = \omega\, p(\theta) + q(\theta) \ ,$$

where p and q are arbitrary functions of θ. Thus (3.6) becomes

$$\frac{\partial \log L}{\partial\theta} = \frac{\partial}{\partial\theta}\sum_{j=1}^{n} \log f(x_j|\theta) = p(\theta)\sum k(x_j) + q(\theta) \ ,$$

whence

$$\frac{\partial}{\partial\theta} \log f(x|\theta) = p(\theta)k(x) + q(\theta)/n \ . \tag{3.9}$$

(3.9) gives the necessary condition for a sufficient statistic to exist. The probability density function is of the form

$$f(x|\theta) = \exp\left\{A(\theta)B(x) + C(x) + D(\theta)\right\} \ . \tag{3.10}$$

Hence under regularity conditions, there is therefore a one-to-one correspondence between the existence of a sufficient statistic for θ and the existence of a MVB estimator of some function of θ.

As a result, there is no sufficient statistic for p.

CHAPTER 4. MAXIMUM LIKELIHOOD ESTIMATOR

Let $x_1, x_2, \ldots, x_m$ be a random sample of size m drawn from a Bounded Above Binomial Distribution. For convenience, let x_i, $i = 1, 2, \ldots, m'$ be such that $x_i < t$ and $x_{m'+1} = x_{m'+2} = \ldots = x_m = t$. The likelihood function is then

$$L(\underline{x},p) = [B(n,p;t^+)]^{m-m'} \left[\prod_{i=1}^{m'} \binom{n}{x_i} \right] p^{\sum_{i=1}^{m'} x_i} (1-p)^{nm' - \sum_{i=1}^{m'} x_i} . \tag{4.1}$$

In particular, if $t = n$, we have $B(n,p;t^+) = Pr(x=n) = p^n$, or if $m' = m$, then (4.1) is reduced to

$$L(\underline{x},p) = \left[\prod_{i=1}^{m} \binom{n}{x_i} \right] p^{\sum_{i=1}^{m} x_i} (1-p)^{nm - \sum_{i=1}^{m} x_i} . \tag{4.2}$$

4.1 Maximum likelihood estimator (MLE)

Let $\hat{\hat{p}}$ be the MLE of p, then it can be shown that $\hat{\hat{p}}$ is the solution of the following equation

$$\hat{\hat{p}} = \frac{\sum_{i=1}^{m'} x_i}{nm'} + \frac{(m-m')t}{nm'} \binom{n}{t} \frac{\hat{\hat{p}}^t (1-\hat{\hat{p}})^{n-t+1}}{B(n,\hat{\hat{p}};t^+)} . \tag{4.3}$$

When $t = n$ (i.e. usual binomial) or when $m' = m$, (4.3) becomes

$$\hat{\hat{p}} = \frac{\sum_{i=1}^{m} x_i}{nm} . \tag{4.4}$$

In the case $m' = 0$, $L(\underline{x},p) = [B(n,p;t^+)]^m$. Since $B(n,p;t^+)$ is a monotone increasing function of p, $L(\underline{x},p)$ is maximum (equal to 1) as $p = 1$. Thus $\hat{\hat{p}} = 1$ when $m' = 0$.

To establish (4.3), it is necessary to express $B(n,p;t^+)$ as the incomplete beta integral (2.2) and employ the following result

$$\frac{\partial}{\partial\theta}\int_a^b f(x)dx = \int_a^b \frac{\partial f(x)}{\partial\theta}dx + f(b)\frac{\partial b}{\partial\theta} - f(a)\frac{\partial a}{\partial\theta} \ , \qquad (4.5)$$

where a, b and f(x) possess continuous first derivatives. Since

$$\log L(\underline{x},p) = (m-m')\log B(n,p;t^+) + \log \prod_{i=1}^{m'} \binom{n}{x_i} + \sum_{i=1}^{m'} x_i \log p$$

$$+ (nm' - \sum_{i=1}^{m'} x_i) \log (1-p) \ ,$$

using (2.2) and (4.5)

$$\frac{\partial \log L(\underline{x},p)}{\partial p} = \frac{(m-m')}{B(n,p;t^+)}\frac{\partial}{\partial p}B(n,p;t^+) + \frac{\sum_{i=1}^{m'} x_i}{p} - \frac{nm' - \sum_{i=1}^{m'} x_i}{1-p}$$

$$= \frac{(m-m')}{B(n,p;t^+)}\frac{\Gamma(n+1)}{\Gamma(t)\Gamma(n-t+1)}(1-p)^{n-t}p^{t-1} + \frac{\sum_{i=1}^{m'} x_i}{p} - \frac{nm' - \sum_{i=1}^{m'} x_i}{1-p} \ .$$

Set $\dfrac{\partial \log L(\underline{x},p)}{\partial p} = 0$ and obtain

$$\frac{(m-m')}{B(n,\hat{\hat{p}};t^+)}\frac{\Gamma(n+1)}{\Gamma(t)\Gamma(n-t+1)}(1-\hat{\hat{p}})^{n-t}\hat{\hat{p}}^{t-1} + \frac{\sum_{i=1}^{m'} x_i}{\hat{\hat{p}}} = \frac{nm' - \sum_{i=1}^{m'} x_i}{1-\hat{\hat{p}}}$$

$$\frac{(m-m')}{B(n,\hat{\hat{p}};t^+)}\frac{\Gamma(n+1)}{\Gamma(t)\Gamma(n-t+1)}(1-\hat{\hat{p}})^{n-t+1}\hat{\hat{p}}^{t} + (1-\hat{\hat{p}})\sum_{i=1}^{m'} x_i = \hat{\hat{p}}\,(nm' - \sum_{i=1}^{m'} x_i)$$

$$\frac{(m-m')t}{B(n,\hat{\hat{p}};t^+)}\frac{n!}{t!(n-t)!}(1-\hat{\hat{p}})^{n-t+1}\hat{\hat{p}}^{t} + \sum_{i=1}^{m'} x_i = \hat{\hat{p}}nm'$$

or

$$\hat{\hat{p}} = \frac{\sum_{i=1}^{m'} x_i}{nm'} + \frac{(m-m')t}{nm'}\binom{n}{t}\frac{\hat{\hat{p}}^t(1-\hat{\hat{p}})^{n-t+1}}{B(n,\hat{\hat{p}};t^+)}$$

It should be note that if no single sufficient statistic for p exists, the likelihood function no longer necessary has a unique maximum value. The problem of uniqueness of (4.3) is now still open. In fact, outside the field of sufficient statistics, the optimum properties of ML estimators are asymptotic ones. We shall see that, under very broad conditions, the most important of which is that the range of $f(x;p)$ does not depend on p, the ML estimator is asymptotically normally distributed and is an efficient estimator.

4.2 The consistency of maximum likelihood estimators

We now proceed to show that the ML estimator is consistent. Before establishing the result, it is necessary to show the existence of $E_o\left\{\frac{1}{m}\log L(\underline{x},p_o)\right\}$, where p_o denotes the true value of p and E_o represents the operation of taking expectation when the true value p_o holds.

Since
$$\log L(\underline{x},p_o) = (m-m')\log B(n,p_o;t^+) + \log\prod_{i=1}^{m'}\binom{n}{x_i} + \sum_{i=1}^{m'} x_i \log p_o + (nm' - \sum_{i=1}^{m'} x_i)\log(1-p_o),$$

hence
$$E_o\left\{\frac{1}{m}\log L(\underline{x},p_o)\right\} = \frac{1}{m}\sum_{\underline{x}} L(\underline{x},p_o)\log L(\underline{x},p_o)$$

$$= \frac{1}{m}\sum_{\underline{x}} L(\underline{x},p_o)\Big[(m-m')\log B(n,p_o;t^+) + \log\prod_{i=1}^{m'}\binom{n}{x_i} + \sum_{i=1}^{m'} x_i\log p_o + (nm' - \sum_{i=1}^{m'} x_i)\log(1-p_o)\Big]$$

$$= \frac{1}{m}\Big\{\log B(n,p_o;t^+)\sum_{\underline{x}} L(\underline{x},p_o)(m-m') + \sum_{\underline{x}} L(\underline{x},p_o)\log\prod_{i=1}^{m'}\binom{n}{x_i}$$

$$+ \log p_o \sum_{\underline{x}} L(\underline{x},p_o) \sum_{i=1}^{m'} x_i + \log(1-p_o) \sum_{\underline{x}} L(\underline{x},p_o)(nm' - \sum_{i=1}^{m'} x_i)$$

$$= \frac{1}{m}\Big\{[\log B(n,p_o;t^+)][m - \sum_{\underline{x}} L(\underline{x},p_o)m'] + \sum_{\underline{x}} L(\underline{x},p_o) \log \prod_{i=1}^{m'} \binom{n}{x_i}$$

$$+ (\log p_o) \sum_{\underline{x}} L(\underline{x},p_o)(\sum_{i=1}^{m'} x_i)$$

$$+ [\log(1-p_o)][n \sum_{\underline{x}} L(\underline{x},p_o)m' - \sum_{\underline{x}} L(\underline{x},p_o)(\sum_{i=1}^{m'} x_i)]\Big\}$$

$$= \frac{1}{m}\Big\{m \log B(n,p_o;t^+) + [\sum_{\underline{x}} L(\underline{x},p_o)m'] \log \frac{(1-p_o)^n}{B(n,p_o;t^+)}$$

$$+ \sum_{\underline{x}} L(\underline{x},p_o) \log \prod_{i=1}^{m'} \binom{n}{x_i} + \sum_{\underline{x}} L(\underline{x},p_o)(\sum_{i=1}^{m'} x_i) \log (\frac{p_o}{1-p_o})\Big\} . \qquad (4.6)$$

For $p_o \in (0,1)$, it is clear that (4.6) exists. Hence $E_o\{\frac{1}{m} \log L(\underline{x},p_o)\}$ exists.

Now we consider the case of m independent observations from a distribution with density $f(x;p)$ and for each m we choose the ML estimator $\hat{\hat{p}}$ so that if p is any admissible value of the parameter, we have

$$\log L(\underline{x},\hat{\hat{p}}) \geq \log L(\underline{x},p) . \qquad (4.7)$$

Consider the random variable $L(\underline{x},p)/L(\underline{x},p_o)$. For all $p^* \neq p_o$, $\sum_{\underline{x}} L(\underline{x},p^*) = \sum_{\underline{x}} L(\underline{x},p_o)$. Recall the proposition (1e.6) of Rao (1973), namely: let $\sum a_i$ and $\sum b_i$ be convergent sequences of positive numbers such that $\sum a_i \geq \sum b_i$, then $\sum a_i \log \frac{b_i}{a_i} \leq 0$, the equality being attained when and only when $a_i = b_i$. Thus we have

$$E_o \left\{ \log \frac{L(\underline{x},p^*)}{L(\underline{x},p_o)} \right\} < 0 , \tag{4.8}$$

or
$$E_o \left\{ \frac{1}{m} \log L(\underline{x},p^*) \right\} < E_o \left\{ \frac{1}{m} \log L(\underline{x},p_o) \right\} , \tag{4.9}$$

as the expectation on the right exists (refer to first paragraph of this section). Now for any value of p,

$$\frac{1}{m} \log L(\underline{x},p) = \frac{1}{m} \sum_{i=1}^{m} \log f(x_i;p)$$

is the mean of a set of m independent identical variates with expectation

$$E_o \{ \log f(\underline{x},p) \} = E_o \left\{ \frac{1}{m} \log L(\underline{x},p) \right\} .$$

By the Strong Law of Large Numbers, $\frac{1}{m} \log L(\underline{x},p)$ converges with probability unity to its expectation, as m increases. Thus for large m, we have, from (4.9) with probability unity

$$\frac{1}{m} \log L(\underline{x},p^*) < \frac{1}{m} \log L(\underline{x},p_o) ,$$

or
$$\lim_{m \to \infty} \text{prob} \left\{ \log L(\underline{x},p^*) < \log L(\underline{x},p_o) \right\} = 1, \quad \text{for all } p^* \neq p_o . \tag{4.10}$$

On the other hand, we obtain from (4.7), with $p = p_o$, that

$$\log L(\underline{x},\hat{\hat{p}}) \geq \log L(\underline{x},p_o) . \tag{4.11}$$

(4.10) and (4.11) implies that

$$\lim_{m \to \infty} \text{prob} \{ \hat{\hat{p}} = p_o \} = 1 ,$$

which establishes the consistency of $\hat{\hat{p}}$.

We are now going to show that there is a unique consistent ML estimator under regularity conditions, as m increases. It is due to Huzurbazar (1948). It is clear that the likelihood function (4.1) possesses two derivatives. Thus following from the convergence in probability of $\hat{\hat{p}}$ to p_o,

$$\frac{1}{m}\left[\frac{\partial^2}{\partial p^2}\log L(\underline{x},p)\right]_{p=\hat{\hat{p}}} \rightarrow \frac{1}{m}\left[\frac{\partial^2}{\partial p^2}\log L(\underline{x},p)\right]_{p=p_o} \tag{4.12}$$

as $m \rightarrow \infty$.

Applying the Strong Law of Large Numbers again,

$$\frac{1}{m}\frac{\partial^2}{\partial p^2}\log L(\underline{x},p) = \frac{1}{m}\sum_{i=1}^{m}\frac{\partial^2}{\partial p^2}\log f(x_i;p)$$

is the mean of m independent identical variates and converges with probability unity to its mean value. Hence (4.12) becomes

$$\lim_{m\rightarrow\infty} \text{prob}\left\{\left[\frac{\partial^2}{\partial p^2}\log L(\underline{x},p)\right]_{p=\hat{\hat{p}}} = E_o\left[\frac{\partial^2}{\partial p^2}\log L(\underline{x},p)\right]_{p=p_o}\right\} = 1 \ . \tag{4.13}$$

Since

$$E\left[\frac{\partial^2}{\partial p^2}\log L(\underline{x},p)\right] = - E\left\{\left(\frac{\partial \log L(\underline{x},p)}{\partial p}\right)^2\right\} < 0 \ ,$$

hence (4.13) becomes

$$\lim_{m\rightarrow\infty} \text{prob}\left\{\left[\frac{\partial^2}{\partial p^2}\log L(\underline{x},p)\right]_{p=\hat{\hat{p}}} < 0\right\} = 1 \ . \tag{4.14}$$

Let $\hat{\hat{p}}_1$ and $\hat{\hat{p}}_2$ be the roots of $(\log L)' = 0$ satisfying $(\log L)'' < 0$. If $\log L(\underline{x},p)$ has a second derivative everywhere, there must be a minimum between the maxima at $\hat{\hat{p}}_1$ and $\hat{\hat{p}}_2$. If this is at $\hat{\hat{p}}_3$, we must have

$$\left[\frac{\partial^2}{\partial p^2}\log L(\underline{x},p)\right]_{p=\hat{\hat{p}}_3} \geq 0 \ . \tag{4.15}$$

Since $\hat{\hat{p}}_1$ and $\hat{\hat{p}}_2$ are consistent estimators, $\hat{\hat{p}}_3$ must be consistent and must satisfy (4.14). However (4.14) and (4.15) directly contradict each other, it follows that we can only have one consistent estimator $\hat{p}$ obtained as a root of the likelihood equation $(\log L)' = 0$. Thus we establish the uniqueness of consistent MLE.

4.3 The efficiency and asymptotic normality of ML estimators

Notice that $L(\underline{x},p)$ possesses continuous first derivative, i.e. $L(\underline{x},p) \in C^1$ and the first two derivatives of $\log L(\underline{x},p)$ also exist. We have

$$E\left(\frac{\partial}{\partial p} \log L(\underline{x},p)\right) = \sum_{\underline{x}} \frac{\partial}{\partial p} \log L(\underline{x},p) \cdot L(\underline{x},p)$$

$$= \sum_{\underline{x}} \frac{1}{L(\underline{x},p)} \frac{\partial}{\partial p} L(\underline{x},p) \cdot L(\underline{x},p)$$

$$= \sum_{\underline{x}} \frac{\partial}{\partial p} L(\underline{x},p)$$

$$= \frac{\partial}{\partial p} \sum_{\underline{x}} L(\underline{x},p) \qquad (\text{since } L(\underline{x},p) \in C^1)$$

$$= 0 \quad , \tag{4.16}$$

as $\sum_{\underline{x}} L(\underline{x},p) = 1$.

Now

$$0 = \frac{\partial}{\partial p} \sum_{\underline{x}} \left(\frac{\partial}{\partial p} \log L(\underline{x},p)\right) \cdot L(\underline{x},p)$$

$$= \sum_{\underline{x}} \left\{ \left[\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\right] L(\underline{x},p) + \frac{\partial}{\partial p} \log L(\underline{x},p) \frac{\partial}{\partial p} L(\underline{x},p) \right\}$$

$$= \sum_{\underline{x}} \left\{ \left[\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\right] L(\underline{x},p) + \left[\frac{\partial}{\partial p} \log L(\underline{x},p)\right]^2 L(\underline{x},p) \right\}$$

$$= E\left(\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\right) + E\left(\frac{\partial}{\partial p} \log L(\underline{x},p)\right)^2 \quad . \tag{4.17}$$

In view of (4.16) and (4.17), we have

$$\begin{aligned} \operatorname{var}\left[\frac{\partial}{\partial p} \log L(\underline{x},p)\right] &= E\left[\left(\frac{\partial}{\partial p} \log L(\underline{x},p)\right)^2\right] \\ &= -E\left[\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\right] \\ &\equiv R^2(p) \ , \end{aligned}$$

where $R^2(p) > 0$.

Using Taylor's theorem for the expansion of $\frac{\partial}{\partial p} \log L(\underline{x},p)\Big|_{\hat{\hat{p}}}$ at p_o, we have

$$\frac{\partial}{\partial p} \log L(\underline{x},p)\Big|_{\hat{\hat{p}}} = \frac{\partial}{\partial p} \log L(\underline{x},p)\Big|_{p_o} + (\hat{\hat{p}}-p_o) \frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\Big|_{p^*} \ , \tag{4.18}$$

where p^* is some value between $\hat{\hat{p}}$ and p_o.

By the fact that $\hat{\hat{p}}$ is a root of the likelihood equation $\frac{\partial}{\partial p} \log L(\underline{x},p) = 0$, we have the left-hand side of (4.18) equals to zero. Thus (4.18) becomes

$$(\hat{\hat{p}} - p_o) = \frac{\frac{\partial}{\partial p} \log L(\underline{x},p)\Big|_{p_o}}{-\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\Big|_{p^*}} \ , \tag{4.19}$$

or

$$\frac{(\hat{\hat{p}} - p_o)}{1/R(p_o)} = \frac{\frac{\partial}{\partial p} \log L(\underline{x},p)\Big|_{p_o} / R(p_o)}{\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\Big|_{p^*} / \left[-R^2(p_o)\right]} \quad . \tag{4.20}$$

Since $\hat{\hat{p}}$ is consistent, thus $\frac{1}{m} \frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\Big|_{p=\hat{\hat{p}}}$ converges to

$\frac{1}{m}\left[\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\right]_{p=p_o}$ in probability. Again, by the Law of Large Number,

$\frac{1}{m}\left[\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\right]_{p=p_o}$ converges in probability to $\frac{1}{m} E_o\left[\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\right]_{p=p_o}$.

Thus

$$\lim_{m\to\infty} \Pr\left\{\left[\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\right]_{p=\hat{\hat{p}}} = E_o\left[\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\right]_{p=p_o}\right\} = 1 .$$

Since p^* lies between $\hat{\hat{p}}$ and p_o,

$$\lim \Pr\left\{\left[\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\right]_{p^*} = - R^2(p_o)\right\} = 1 , \tag{4.21}$$

so that the denominator converges to unity.

For the numerator, since $\frac{\partial}{\partial p} \log L(\underline{x},p)$ is the sum of m independent identical variables each with expectation zero and variance $\frac{1}{m} R^2(p_o)$, thus by applying the Central Limit Theorem, we then have

$$\frac{\partial}{\partial p} \log L(\underline{x},p)\Big|_{p_o} \mathrel{\dot\sim} N(0, R^2(p_o)) ,$$

therefore
$$\left[\frac{\partial}{\partial p} \log L(\underline{x},p)\Big|_{p_o} - 0\right] / R(p_o) \mathrel{\dot\sim} N(0, 1) . \tag{4.22}$$

Hence
$$\frac{(\hat{\hat{p}} - p_o)}{1/R(p_o)} = \frac{\frac{\partial}{\partial p} \log L(\underline{x},p)\Big|_{p_o}/R(p_o)}{\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\Big|_{p^*}/\left[- R^2(p_o)\right]} \mathrel{\dot\sim} N(0, 1) . \tag{4.23}$$

Thus $\hat{\hat{p}} \mathrel{\dot\sim} N(p_o, 1/R^2(p_o))$.

In other words, the ML estimator $\hat{\hat{p}}$ is asymptotically normally distributed.

This result gives the ML estimator an asymptotic variance equal to the Cramér-Rao MVB and thus implies that the ML estimator is efficient.

Now, we have

$$\operatorname{var}(\hat{\hat{p}}) \doteq \frac{1}{R^2(p_o)} = -1 \Big/ E_o\left[\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\right]_{p=p_o}$$

$$= -1 \Big/ mE_o\left[\frac{\partial^2}{\partial p^2} \log f(x;p)\right]_{p=p_o}, \qquad (4.24)$$

where m is the sample size, and

$$E_o\left[\frac{\partial^2}{\partial p^2} \log L(\underline{x},p)\right]_{p=p_o} = \sum_{\underline{x}} L(\underline{x},p_o)\frac{\partial^2}{\partial p^2} \log L(\underline{x},p_o)$$

$$= \frac{t\binom{n}{t}}{B(n,p_o;t^+)} p_o^{t-2}(1-p_o)^{n-t-1}\left[(t-1+p_o-p_on) - \frac{t\binom{n}{t} p_o^t(1-p_o)^{n-t+1}}{B(n,p_o;t^+)}\right]$$

$$\cdot\left[m - \sum_{\underline{x}} L(\underline{x},p_o)m'\right] + \left[\frac{1}{(1-p_o)^2} - \frac{1}{p_o^2}\right] \sum_{\underline{x}} L(\underline{x},p_o)\left(\sum_{i=1}^{m'} x_i\right)$$

$$- \frac{n}{(1-p_o)^2} \sum_{\underline{x}} L(\underline{x},p_o)m' .$$

Thus

$$\operatorname{var}(\hat{\hat{p}}) \doteq -\left\{\frac{t\binom{n}{t}}{B(n,p_o;t^+)} p_o^{t-2}(1-p_o)^{n-t-1}\left[(t-1+p_o-p_on) - \frac{t\binom{n}{t} p_o^t(1-p_o)^{n-t+1}}{B(n,p_o;t^+)}\right]\right.$$

$$\cdot\left[m - \sum_{\underline{x}} L(\underline{x},p_o)m'\right] + \left[\frac{1}{(1-p_o)^2} - \frac{1}{p_o^2}\right] \sum_{\underline{x}} L(\underline{x},p_o)\left(\sum_{i=1}^{m'} x_i\right)$$

$$\left.- \frac{n}{(1-p_o)^2} \sum_{\underline{x}} L(\underline{x},p_o)m'\right\}^{-1} . \qquad (4.25)$$

Notice that (4.25) is fairly complicated, thus we seek for a simpler expression which enable us to calculate var($\hat{\hat{p}}$), say, by using a desk calculator and appropriate Binomial table.

To express (4.25) in another form, we first notice that

$$\frac{\partial^2}{\partial p^2} \log f(x;p) = \begin{cases} -\frac{x}{p^2} - \frac{n-x}{(1-p)^2}, & \text{if } 0 \leq x < t ; \\ \frac{t\binom{n}{t} p^{t-2}(1-p)^{n-t-1}}{B(n,p;t^+)} \left[(t-1+p-pn) - \frac{t\binom{n}{t} p^t(1-p)^{n-t+1}}{B(n,p;t^+)} \right], & \text{if } x = t . \end{cases}$$

Hence

$$E\left[\frac{\partial^2}{\partial p^2} \log f(x;p)\right] = \sum_{x=o}^{t} \frac{\partial^2}{\partial p^2} \log f(x;p)\cdot f(x;p)$$

$$= \sum_{x=o}^{t-1} \binom{n}{x} p^x(1-p)^{n-x} \left[-\frac{x}{p^2} - \frac{n-x}{(1-p)^2}\right] + t\binom{n}{t} p^{t-2}(1-p)^{n-t-1}$$

$$\cdot \left[(t-1+p-pn) - \frac{t\binom{n}{t} p^t(1-p)^{n-t+1}}{B(n,p;t^+)}\right]$$

$$= \sum_{x=o}^{t-1} \binom{n}{x} p^x(1-p)^{n-x} \left[x\frac{2p-1}{p^2(1-p)^2} - \frac{n}{(1-p^2)}\right] + t\binom{n}{t} p^{t-2}(1-p)^{n-t-1}$$

$$\cdot \left[(t-1+p-pn) - \frac{t\binom{n}{t} p^t(1-p)^{n-t+1}}{B(n,p;t^+)}\right] . \qquad (4.26)$$

Since

$$\sum_{x=o}^{t-1} \binom{n}{x} p^x(1-p)^{n-x}\cdot x = \sum_{x=1}^{t-1} \binom{n}{x} p^x(1-p)^{n-x}\cdot x$$

$$= \sum_{x=1}^{t-1} \frac{n(n-1)!}{(x-1)!(n-x)!} p^x(1-p)^{n-x}$$

$$= np \sum_{x=1}^{t-1} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1}(1-p)^{n-x}$$

$$= np \sum_{x=1}^{t-1} \binom{n-1}{x-1} p^{x-1}(1-p)^{n-x}$$

$$= np \sum_{x=0}^{t-2} \binom{n-1}{x} p^{x}(1-p)^{n-1-x}$$

$$= np\, B[n-1,p;(t-2)^{-}] ,$$

thus (4.26) can be written as

$$E\left[\frac{\partial^2}{\partial p^2} \log f(x;p)\right] = \frac{(2p-1)n}{p(1-p)^2} B[n-1,p;(t-2)^{-}] - \frac{n}{(1-p)^2} B[n,p;(t-1)^{-}]$$

$$+ t\binom{n}{t} p^{t-2}(1-p)^{n-t-1}\left[(t-1+p-pn) - \frac{t\binom{n}{t} p^{t}(1-p)^{n-t+1}}{B(n,p;t^{+})}\right], \qquad (4.27)$$

and

$$\operatorname{var}(\hat{p}) \doteq -\frac{1}{m} \frac{1}{E_o\left[\frac{\partial^2}{\partial p^2} \log f(x;p)\right]_{p=p_o}} ,$$

which is much simpler than (4.25).

4.4 Successive approximation to ML estimator

The MLE obtained from (4.3) is in implicit form and is so complicated that iterative methods must be used. However, successive approximation to the MLE may be obtained by applying the Newton's method (Hildebrand (1956)).

Let c be another consistent estimator of p, then

$$0 = \left(\frac{\partial \log L}{\partial p}\right)_{\hat{p}} = \left(\frac{\partial \log L}{\partial p}\right)_{c} + (\hat{p}-c)\left(\frac{\partial^2 \log L}{\partial p^2}\right)_{p^*} ,$$

where p^* lies between $\hat{p}$ and c.

Since $\hat{\hat{p}}$ and c are both consistent, therefore, in large sample, p* converges to p_o in probability. Thus $\left(\frac{\partial^2 \log L}{\partial p^2}\right)_{p*}$ converges in probability to $\left(\frac{\partial^2 \log L}{\partial p^2}\right)_{p_o}$ which is $-1/\mathrm{var}(\hat{\hat{p}})\big|_{p=p_o}$. Hence

$$\hat{\hat{p}} = c + \left(\frac{\partial \log L}{\partial p}\right)_c \cdot \mathrm{var}(\hat{\hat{p}})\big|_{p=p_o} . \tag{4.28}$$

With $\mathrm{var}(\hat{\hat{p}})$ estimated from the sample if necessary, (4.28) will give a closer approximation to a root of the likelihood equation. The operation can be repeated, carrying on if necessary until no further correction is achieved.

It should be noted that there is no guarantee that the root of the likelihood equation obtained in this way will correspond to the absolute maximum of the likelihood function. However, for large m, the fact (Section 4.2) that there is a unique consistent root comes to our aid.

CHAPTER 5. MOMENT ESTIMATOR AND OTHER ESTIMATORS

5.1 Moment estimator

Consider a random sample $x_1, x_2, \ldots, x_m$ drawn from a Bounded Above Binomial Distribution with parameters n, p and t. Let M_r' denotes the rth noncentral sample moment, that is,

$$M_r' = \sum_{i=1}^{m} x_i^r/m \quad . \tag{5.1}$$

Then the moment estimator, $\hat{p}$, of p is obtained by setting M_1' equal to the population first noncentral moment μ_1' . Thus

$$\sum_{i=1}^{m} x_i/m = n\hat{p}\ B\left[n-1,\hat{p};(t-2)^-\right] + t\ B(n,\hat{p};t^+) \ , \tag{5.2}$$

or in the iterative form

$$\hat{p}_{k+1} = \frac{\overline{x} - t\ B(n,\hat{p}_k;t^+)}{n\ B\left[n-1,\hat{p}_k;(t-2)^-\right]} \ , \tag{5.3}$$

in which $\overline{x}$ is M_1' and $\hat{p}_0$ is a initial guess.

Notice that for a fixed n, a smaller p will yield a smaller observation x. Thus, if all the observed sample items are less than t, i.e. m' = m, then $p \doteq 0$ will be a reasonable estimate. Recall that $B(n,p;t^+)$ is a monotone increasing function of p. If we put $\hat{p}_0 = 0$ on the right hand side of (5.3), then

$$\hat{p} = \frac{\overline{x}}{n}$$

i.e.
$$\hat{p} = \sum_{i=1}^{m} x_i/nm ,$$

identical to the MLE in the case m' = m.

5.2 Consistency of the moment estimator

Observe that M_r' is the mean of m independent identical random variables, namely $M_r' = \sum_{i=1}^{m} Y_i/m$ with $Y_i = x_i^r$, and from (2.10) and (2.11), $E(Y_i) = \mu_r'$ exists. Thus by the Law of Large Numbers, M_r' converges to μ_r' with probability unity so that M_r' is a consistent estimator of μ_r'.

Now
$$\mu_1' = np\, B[n-1,p;(t-2)^-] + t\, B(n,p;t^+)$$
$$= np \sum_{j=o}^{t-2} \binom{n-1}{j} p^j(1-p)^{n-1-j} + t \sum_{j=t}^{n} \binom{n}{j} p^j(1-p)^{n-j}$$

is a continuous function of p (in fact is a polynomial in p). If p_o denotes the true (unknown) value of p, then in an interval of p_o, μ_1' is one to one. Hence the inverse function (implicit) of p exists in this interval and is continuous, i.e. $p = h(\mu_1')$.

From the continuity, we note that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that if
$$|M_1' - \mu_1'| < \delta \quad , \tag{5.4}$$
then
$$|h(M_1') - h(\mu_1')| < \varepsilon \quad . \tag{5.5}$$

That is, (5.4) is a subset of the (5.5); thus

$$\Pr\left[|M_1' - \mu_1'| < \delta\right] \leq \Pr\left[|h(M_1') - h(\mu_1')| < \varepsilon\right] .$$

However, consistency of M_1' means that

$$\lim_{m \to \infty} \Pr\left[|M_1' - \mu_1'| < \delta\right] = 1 . \qquad (5.6)$$

Since probability is less than or equal to one, it follows that

$$\lim_{m \to \infty} \Pr\left[|h(M_1') - h(\mu_1')| < \varepsilon\right] = 1 , \qquad (5.7)$$

and hence, $\hat{p} = h(M_1')$, implicit solution of (5.2), is a consistent estimator of $h(\mu_1') = p$.

5.3 Other estimators

5.3.1 Pooled estimate

Recall that in the usual binomial case, the maximum likelihood estimate and the moment estimate of p are identical and equal to x/n which is unbiased, consistent and sufficient. Thus, regardless of the "bounded" configuration,

$$\check{p} = \sum_{i=1}^{m} x_i/nm \qquad (5.8)$$

may serve as a pooled estimate of p based on a random sample of size m drawn from a Bounded Above Binomial Distribution with parameters n, p and t. However, $\check{p}$ is not unbiased except when t = n. In fact,

$$E(\check{p}) = p\, B\left[n-1, p; (t-2)^-\right] + \frac{t}{n} B(n, p; t^+) .$$

It is obvious that $\check{p}$ underestimates p whenever some observed x_i's attain the bound.

5.3.2 Selected estimate

Another estimator more underestimating p will be

$$\tilde{p} = \sum_{i=1}^{m'} x_i/nm' \quad , \tag{5.9}$$

where, as before, $x_1, x_2, \ldots, x_{m'}$ denote those observed sample items which are less than t.

5.3.3 Ratio estimate

Observe that $\Pr(x < t | n,p,t) = B[n,p;(t-1)^-]$. If in a random sample of m items, m' of them are less than t, then the ratio m'/m may serve as an estimate of $\Pr(x < t)$. Thus we define the ratio estimate $\bar{p}$ as the solution of the equation

$$\frac{m'}{m} = B[n,\bar{p};(t-1)^-] \ . \tag{5.10}$$

When $m' = 0$ (i.e. $x_i = t$ for all i), $\bar{p} = 1$ and when $m' = m$ (i.e. $x_i < t$ for all i), $\bar{p} = 0$. It is clear that both estimates may be too extreme, that is, overestimating and underestimating the true value of p respectively. A graphical solution of $\bar{p}$ in (5.10) may easily be obtained from Larson (1966)'s nomograph, which relates each of n, p, c, P_a to the others, where

$$P_a = \sum_{j=0}^{c} \binom{n}{j} p^j (1-p)^{n-j} \ .$$

CHAPTER 6. COMPARISON OF THE VARIOUS ESTIMATES OF THE PARAMETER p BY USING COMPUTER SIMULATION

In the previous chapters, we have already establish various estimates of p for the Bounded Above Binomial Distribution. Now we are going to the discussion of comparison of various estimates.

We carry out this job by using computer simulation which consists of the following steps:

(1) Define p_o (true value of p), $0 < p_o < 1$,

n and t such that $t < n$, and the sample size m .

(2) (i) Generate uniform random numbers u_j $(u_j \in (0,1))$, $j = 1, 2, \dots , n$.

If $u_j \leq p_o$, let $y_j = 1$, otherwise let $y_j = 0$.

(ii) Let

$$x_i = \sum_{j=1}^{n} y_j .$$

If $x_i \geq t$, put $x_i = t$.

Step (2) is repeated for m times to obtain m x_i's, namely $x_1, x_2, \dots , x_m$, which constitutes a random sample of size m from a Bounded Above Binomial Distribution with parameters n, p and t.

(3) Obtain the various estimates.

For convenience, denote those sample items which are less than t, m' such items say, as $x_1, x_2, \dots , x_{m'}$. Then obtain the following estimates:

(a) Obtain the ML estimate $\hat{\hat{p}}$ by using the iteration formula (4.3)

$$\hat{\hat{p}}_{k+1} = \frac{\sum_{i=1}^{m'} x_i}{nm'} + \frac{(m-m')t}{nm'} \binom{n}{t} \frac{\hat{\hat{p}}_k^t (1-\hat{\hat{p}}_k)^{n-t+1}}{B(n,\hat{\hat{p}}_k;t^+)} . \qquad (6.1)$$

The initial estimates $\hat{\hat{p}}_o$ of (6.1) can be chosen arbitrary so long as $\hat{\hat{p}}_o \in (0,1)$. The iteration is stopped if $|\hat{\hat{p}}_{k+1} - \hat{\hat{p}}_k| < 10^{-6}$.

(b) Obtain the moment estimate $\hat{p}$ by using the iteration formula (5.3)

$$\hat{p}_{k+1} = \frac{\bar{x} - t\ B(n,\hat{p}_k;t^+)}{n\ B[n-1,\hat{p}_k;(t-2)^-]} , \qquad (6.2)$$

where $\bar{x} = \dfrac{\sum_{i=1}^{m} x_i}{m}$.

Again the initial estimates $\hat{p}_o$ of (6.2) can be chosen arbitrary so long as $\hat{p}_o \in (0,1)$. The iteration is allowed to stop if $|\hat{p}_{k+1} - \hat{p}_k| < 10^{-6}$.

(c) The pooled estimate $\check{p}$ is obtained from (5.8)

$$\check{p} = \frac{\sum_{i=1}^{m} x_i}{nm} . \qquad (6.3)$$

(d) The selected estimate $\tilde{p}$ is obtained from (5.9)

$$\tilde{p} = \frac{\sum_{i=1}^{m'} x_i}{nm'} . \qquad (6.4)$$

(e) The ratio estimate $\bar{p}$ is obtained from (5.10)

$$\frac{m'}{m} = B[n,\bar{p};(t-1)^-] \qquad (6.5)$$

As pointed out in previous chapter, $\bar{p}$ can be obtained graphically from Larson's nomograph. For $\hat{\hat{p}}$ and $\hat{p}$, the IMSL routine ZXSSQ is called for in performing the iteration procedure. This subroutine is based on a finite difference Levenberg-Marquardt algorithm.

(4) For each set of values of p_o, n, t, and m, repeat the whole procedure, step (1) through step (3), for r = 20 times, and record m_i, $\hat{\hat{p}}_i$, $\hat{p}_i$, $\check{p}_i$, $\tilde{p}_i$ and $\bar{p}_i$, i = 1, 2, ... , r. Finally evaluate the mean and variance of the various estimates.

Since it is impossible to reproduce all the output for various values of n, t, p and m, only some concrete examples are given. In this simulation study, we choose m = 25(25)100 for each p_o with different values of n and t where p_o = 0.1(.1).9 . By using the computer, IBM-370/3031 at The Chinese University of Hong Kong, we get the following numerical examples. Each example consists of 20 data sets but only one of them is given. Tables for the mean and variance for various p_i's are also given but with the individual p_i's omitted.

Example 1) $p_o = .1$, $n = 20$, $t = 3$, $m = 25$

x_i:	0 3 1 3 3 2 2 3 1 3	$m' = 15$
	2 3 2 3 1 1 3 3 2 1	$\hat{\hat{p}} = .111669$ with $k = 5$ iterations
	2 3 0 1 1	$\hat{p} = .111615$ with $k = 7$ iterations
		$\check{p} = .098000$
		$\tilde{p} = .063333$
		$\bar{p} = .113472$

Example 2) $p_o = .1$, $n = 20$, $t = 3$, $m = 50$

x_i:	2 3 2 3 1 0 2 2 2 1	$m' = 31$
	3 1 3 1 0 1 2 1 2 2	$\hat{\hat{p}} = .108483$ with $k = 6$ iterations
	3 3 2 3 1 3 2 2 3 3	$\hat{p} = .108064$ with $k = 9$ iterations
	1 3 3 0 3 0 3 0 3 3	$\check{p} = .096000$
	0 3 3 1 1 2 2 1 2 3	$\tilde{p} = .062903$
		$\bar{p} = .109961$

Example 3) $p_o = .1$, $n = 20$, $t = 3$, $m = 75$

x_i:	1 2 3 3 1 1 0 2 3 3	$m' = 53$
	2 3 2 1 3 0 3 2 2 1	$\hat{\hat{p}} = .096205$ with $k = 6$ iterations
	1 1 3 3 2 1 2 1 2 0	$\hat{p} = .096520$ with $k = 9$ iterations
	3 3 0 3 2 2 2 2 1 1	$\check{p} = .088000$
	1 1 3 3 2 1 1 3 3 3	$\tilde{p} = .062264$
	1 2 1 1 2 0 1 3 3 0	$\bar{p} = .094765$
	2 1 1 2 2 1 3 1 3 3	
	0 1 1 1 2	

Example 4) $p_o = .1$, $n = 20$, $t = 3$, $m = 100$

x_j:

0	0	2	1	2	3	2	2	2	1
1	0	2	1	3	0	2	0	2	1
3	2	2	3	2	3	2	1	3	1
3	1	0	0	2	3	1	1	1	3
3	3	2	0	0	1	1	3	1	2
3	3	2	1	1	2	2	2	3	3
2	3	2	1	2	3	0	2	0	3
3	2	2	2	1	3	1	1	1	1
2	2	1	1	3	0	2	2	1	0
2	3	0	3	2	3	0	0	1	2

$m' = 75$

$\hat{\hat{p}} = .089899$ with $k = 5$ iterations

$\hat{p} = .090441$ with $k = 9$ iterations

$\check{p} = .083500$

$\tilde{p} = .061333$

$\bar{p} = .099463$

Example 5) $p_o = .2$, $n = 30$, $t = 7$, $m = 25$

x_i:

7	7	7	7	3	4	4	7	6	4
7	3	3	5	7	6	7	7	3	7
5	7	4	7	6					

$m' = 13$

$\hat{\hat{p}} = .207026$ with $k = 3$ iterations

$\hat{p} = .203608$ with $k = 5$ iterations

$\check{p} = .186667$

$\tilde{p} = .143589$

$\bar{p} = .216118$

Example 6) $p_o = .2$, $n = 30$, $t = 7$, $m = 50$

x_i:

6	4	6	7	5	4	5	7	7	7
6	5	7	5	4	6	6	7	5	7
6	7	7	7	6	4	2	7	6	6
7	7	5	7	5	1	7	7	7	7
7	7	2	4	7	6	7	6	7	7

$m' = 26$

$\hat{\hat{p}} = .218531$ with $k = 3$ iterations

$\hat{p} = .219626$ with $k = 5$ iterations

$\check{p} = .196000$

$\tilde{p} = .161538$

$\bar{p} = .216118$

Example 7) $p_c = .2$, $n = 30$, $t = 7$, $m = 75$

x_i:	6	6	7	7	7	7	4	7	7	5	$m' = 36$
	7	5	6	7	7	6	2	7	5	7	$\hat{\hat{p}} = .224910$ with $k = 3$ iterations
	7	5	7	7	7	5	5	7	5	6	$\hat{p} = .225554$ with $k = 5$ iterations
	5	7	0	7	6	6	7	4	4	7	$\check{p} = .199111$
	7	5	6	7	6	7	5	7	7	6	$\tilde{p} = .162037$
	7	7	2	7	7	7	7	7	4	7	$\bar{p} = .223634$
	4	6	6	5	4	5	7	7	7	7	
	4	7	6	5	7						

Example 8) $p_c = .2$, $n = 30$, $t = 7$, $m = 100$

x_i:	7	6	6	3	7	6	7	6	2	5	$m' = 57$
	4	6	7	6	5	5	2	7	7	7	$\hat{\hat{p}} = .212468$ with $k = 3$ iterations
	7	5	7	6	5	4	5	5	7	7	$\hat{p} = .214804$ with $k = 5$ iterations
	6	7	6	5	5	4	7	5	6	6	$\check{p} = .193333$
	7	6	3	7	4	7	5	3	5	7	$\tilde{p} = .163158$
	7	5	6	7	3	5	2	6	7	7	$\bar{p} = .206848$
	7	6	5	6	7	7	7	3	7	7	
	7	6	7	3	3	6	6	7	7	5	
	5	7	4	7	5	6	7	7	7	7	
	7	7	7	6	7	6	7	4	6	7	

Example 9) $p_o = .3$, $n = 40$, $t = 15$, $m = 25$

x_i: 12 15 10 10 15 11 11 13 14 9
10 15 15 10 9 11 15 14 9 12
14 15 7 10 15

$m' = 18$

$\hat{\hat{p}} = .310249$ with $k = 6$ iterations

$\hat{p} = .307557$ with $k = 2$ iterations

$\check{p} = .301000$

$\tilde{p} = .272222$

$\bar{p} = .320459$

Example 10) $p_o = .3$, $n = 40$, $t = 15$, $m = 50$

x_i: 11 8 15 11 10 15 15 11 15 9
11 14 9 13 9 9 15 11 13 8
9 14 7 11 11 14 11 13 15 12
12 15 15 8 15 8 10 9 8 15
11 14 6 15 13 13 12 8 8 15

$m' = 38$

$\hat{\hat{p}} = .296629$ with $k = 2$ iterations

$\hat{p} = .294202$ with $k = 2$ iterations

$\check{p} = .289500$

$\tilde{p} = .262500$

$\bar{p} = .311554$

Example 11) $p_o = .3$, $n = 40$, $t = 15$, $m = 75$

x_i: 13 12 15 15 8 13 5 13 9 15
14 10 13 7 10 15 10 12 15 11
10 9 9 11 15 15 9 11 11 12
9 15 13 15 9 15 12 13 13 11
13 13 9 9 12 12 11 14 9 15
15 7 15 10 13 12 12 9 9 6
11 15 13 12 13 12 15 11 13 9
13 12 11 9 13

$m' = 60$

$\hat{\hat{p}} = .299052$ with $k = 2$ iterations

$\hat{p} = .298214$ with $k = 2$ iterations

$\check{p} = .293000$

$\tilde{p} = .272500$

$\bar{p} = .301905$

Example 12) $p_o = .3$, $n = 40$, $t = 15$, $m = 100$

x_i:

10	9	10	11	13	10	13	15	12	13
10	15	10	15	15	11	9	11	14	15
11	13	14	15	11	12	10	8	13	12
12	8	10	13	13	9	11	12	9	15
15	10	13	15	15	5	15	15	9	14
14	14	13	12	15	15	15	7	13	10
9	9	10	13	9	15	9	10	10	13
14	14	11	8	10	10	15	15	8	11
15	9	15	15	8	15	12	11	11	12
13	13	12	12	12	15	14	13	15	11

$m' = 76$

$\hat{\hat{p}} = .308294$ with $k = 2$ iterations

$\hat{p} = .306965$ with $k = 2$ iterations

$\check{p} = .300500$

$\tilde{p} = .276974$

$\bar{p} = .311554$

Example 13) $p_o = .4$, $n = 50$, $t = 20$, $m = 25$

x_i:

19	19	20	19	16	20	20	20	19	20
19	17	20	19	16	20	20	20	20	20
20	16	19	20	20					

$m' = 11$

$\hat{\hat{p}} = .413415$ with $k = 4$ iterations

$\hat{p} = .423354$ with $k = 5$ iterations

$\check{p} = .382400$

$\tilde{p} = .360000$

$\bar{p} = .401141$

Example 14) $p_o = .4$, $n = 50$, $t = 20$, $m = 50$

x_i:

19	20	20	18	19	20	20	16	20	20
20	20	14	20	19	20	20	20	20	16
13	20	17	19	20	19	17	20	20	20
14	20	20	20	20	17	20	14	18	18
18	20	17	16	20	20	20	19	19	20

$m' = 22$

$\hat{\hat{p}} = .402872$ with $k = 4$ iterations

$\hat{p} = .403955$ with $k = 5$ iterations

$\check{p} = .374400$

$\tilde{p} = .341818$

$\bar{p} = .401141$

Example 15) $p_o = .4$, $n = 50$, $t = 20$, $m = 75$

x_i:

20	20	20	17	20	16	20	20	18	18
20	19	16	15	20	14	19	15	17	18
17	20	20	20	18	20	20	20	17	20
20	16	17	17	19	19	19	20	19	13
20	19	20	20	20	18	14	17	20	14
14	20	18	20	20	20	20	20	20	16
18	17	19	20	20	20	20	20	17	20
20	20	20	20	20					

$m' = 35$

$\hat{\hat{p}} = .397744$ with $k = 6$ iterations

$\hat{p} = .398453$ with $k = 5$ iterations

$\check{p} = .371733$

$\tilde{p} = .339429$

$\bar{p} = .396498$

Example 16) $p_o = .4$, $n = 50$, $t = 20$, $m = 100$

x_i:

20	18	17	19	12	20	20	18	19	18
20	20	18	19	18	18	20	14	20	18
20	20	20	20	17	20	20	20	18	20
20	20	20	20	20	20	17	20	20	20
20	20	20	20	20	20	20	15	18	20
17	20	16	18	20	20	16	19	19	20
19	20	19	20	20	20	20	17	14	20
20	20	20	20	20	20	19	17	17	20
16	17	20	16	16	20	17	15	20	20
20	13	20	17	20	20	20	18	20	20

$m' = 40$

$\hat{\hat{p}} = .408543$ with $k = 4$ iterations

$\hat{p} = .408795$ with $k = 5$ iterations

$\check{p} = .376600$

$\tilde{p} = .341500$

$\bar{p} = .408234$

Example 17) $p_o = .5$, $n = 60$, $t = 30$, $m = 25$

x_i:

23	30	26	30	26	30	27	30	30	30
30	30	30	28	30	22	29	30	29	26
30	21	24	26	30					

$m' = 12$

$\hat{\hat{p}} = .486253$ with $k = 4$ iterations

$\hat{p} = .482499$ with $k = 5$ iterations

$\check{p} = .464667$

$\tilde{p} = .426389$

$\bar{p} = .494940$

Example 18) $p_o = .5$, $n = 60$, $t = 30$, $m = 50$

x_i:									
26	28	24	28	29	29	30	21	29	30
23	23	30	29	29	26	30	30	23	30
29	30	30	30	23	25	30	25	30	30
30	25	30	29	28	27	30	30	30	30
22	21	30	29	30	24	30	29	30	26

$m' = 28$

$\hat{\hat{p}} = .480411$ with $k = 6$ iterations

$\hat{p} = .479786$ with $k = 5$ iterations

$\check{p} = .463000$

$\tilde{p} = .433929$

$\bar{p} = .482009$

Example 19) $p_o = .5$, $n = 60$, $t = 30$, $m = 75$

x_i:									
24	27	29	30	29	30	30	30	25	30
25	25	30	30	30	30	29	29	30	29
27	30	26	30	30	29	28	30	30	30
27	28	29	23	30	30	30	29	30	30
30	23	30	26	30	30	30	30	30	26
30	26	28	28	30	24	25	30	28	30
30	23	24	23	27	28	30	27	26	30
28	30	30	30	30					

$m' = 36$

$\hat{\hat{p}} = .496079$ with $k = 4$ iterations

$\hat{p} = .496696$ with $k = 5$ iterations

$\check{p} = .472667$

$\tilde{p} = .443056$

$\bar{p} = .494940$

Example 20) $p_o = .5$, $n = 60$, $t = 30$, $m = 100$

x_i:

30	28	30	30	29	30	30	30	24	30
30	30	29	30	29	30	30	30	30	28
30	29	27	30	23	30	30	30	30	30
25	26	30	30	30	26	23	26	30	30
30	28	30	30	24	29	28	30	29	30
30	30	29	30	29	30	26	29	30	30
30	30	24	25	19	26	26	30	30	21
30	30	30	29	30	30	29	30	29	27
30	29	27	29	30	30	30	27	30	26
29	30	26	21	30	29	30	30	24	30

$m' = 43$

$\hat{\hat{p}} = .503026$ with $k = 4$ iterations

$\hat{p} = .503021$ with $k = 5$ iterations

$\check{p} = .475833$

$\tilde{p} = .443798$

$\bar{p} = .503064$

Example 21) $p_o = .6$, $n = 65$, $t = 40$, $m = 25$

x_i:

35	40	38	40	40	40	38	40	33	35
38	37	39	40	38	37	40	40	40	39
36	39	38	38	40					

$m' = 15$

$\hat{\hat{p}} = .605414$ with $k = 6$ iterations

$\hat{p} = .611784$ with $k = 5$ iterations

$\check{p} = .589538$

$\tilde{p} = .572308$

$\bar{p} = .591764$

Example 22) $p_o = .6$, $n = 65$, $t = 40$, $m = 50$

x_i:	40	40	33	34	40	39	32	40	40	40	$m' = 24$
	40	37	39	40	40	40	34	40	35	40	$\hat{\hat{p}} = .602987$ with $k = 4$ iterations
	32	40	40	38	34	40	40	37	39	39	
	34	40	40	34	39	40	34	38	39	40	$\hat{p} = .599907$ with $k = 5$ iterations
	40	38	40	40	34	40	40	34	28	40	

$\check{p} = .582769$

$\tilde{p} = .547436$

$\bar{p} = .610162$

Example 23) $p_o = .6$, $n = 65$, $t = 40$, $m = 75$

x_i:	40	39	39	40	34	37	40	40	32	37	$m' = 42$
	33	40	39	36	34	40	40	36	36	40	$\hat{\hat{p}} = .596980$ with $k = 6$ iterations
	39	38	36	34	40	40	32	40	38	39	
	40	40	35	40	40	37	31	39	39	40	$\hat{p} = .596589$ with $k = 5$ iterations
	40	35	36	38	36	40	40	40	40	40	
	34	40	39	34	33	33	40	40	40	40	
	35	40	40	30	36	40	40	33	40	37	
	39	37	40	38	39						

$\check{p} = .580718$

$\tilde{p} = .553480$

$\bar{p} = .597991$

Example 24) $p_o = .6$, $n = 65$, $t = 40$, $m = 100$

x_i:

35	35	29	40	36	35	33	36	39	40
37	39	36	37	40	40	40	36	40	40
39	36	36	40	33	35	39	40	37	34
40	39	39	31	40	40	40	37	33	33
40	40	33	40	34	40	40	39	37	36
40	35	40	40	36	39	37	40	40	39
35	32	37	34	40	39	40	37	39	32
40	40	40	40	40	40	34	40	36	36
37	40	40	34	40	40	38	36	40	37
38	36	40	35	40	35	40	40	40	37

$m' = 58$

$\hat{\hat{p}} = .594074$ with $k = 6$ iterations

$\hat{p} = .593772$ with $k = 5$ iterations

$\check{p} = .578923$

$\tilde{p} = .552519$

$\bar{p} = .590461$

Example 25) $p_o = .7$, $n = 70$, $t = 50$, $m = 25$

x_i:

47	50	50	50	43	49	50	48	50	50
50	50	47	50	47	50	47	48	50	45
50	50	49	49	50					

$m' = 11$

$\hat{\hat{p}} = .719001$ with $k = 4$ iterations

$\hat{p} = .721941$ with $k = 5$ iterations

$\check{p} = .696571$

$\tilde{p} = .674026$

$\bar{p} = .714302$

Example 26) $p_o = .7$, $n = 70$, $t = 50$, $m = 50$

x_i:

48	44	48	50	50	48	49	50	47	50
50	48	47	50	48	49	49	47	48	45
50	48	45	43	45	50	47	50	46	50
44	48	47	50	45	50	50	48	50	50
50	50	48	50	47	47	42	48	50	48

$m' = 31$

$\hat{\hat{p}} = .698574$ with $k = 6$ iterations

$\hat{p} = .702053$ with $k = 5$ iterations

$\check{p} = .686000$

$\tilde{p} = .668664$

$\bar{p} = .689381$

Example 27) $p_o = .7$, $n = 70$, $t = 50$, $m = 75$

x_i:

50	45	46	50	47	43	50	50	49	44
42	48	47	44	50	49	44	50	50	40
48	50	50	49	50	42	47	45	49	47
48	47	50	50	49	49	38	46	50	50
50	45	43	50	47	48	49	50	50	50
50	47	46	50	49	50	47	49	46	44
46	50	49	50	50	50	50	49	44	44
50	44	49	50	49					

$m' = 46$

$\hat{\hat{p}} = .693346$ with $k = 6$ iterations

$\hat{p} = .694340$ with $k = 5$ iterations

$\check{p} = .681143$

$\tilde{p} = .660248$

$\bar{p} = .690344$

Example 28) $p_o = .7$, $n = 70$, $t = 50$, $m = 100$

x_i:	49	50	48	50	50	49	50	47	50	44
	46	48	50	44	47	48	50	50	49	50
	50	48	50	50	50	46	46	49	46	47
	50	50	49	50	47	46	48	50	50	50
	44	42	49	47	50	43	50	50	50	50
	48	48	50	46	50	49	49	50	50	50
	49	49	50	50	50	43	50	48	44	45
	48	43	50	46	50	40	45	47	41	50
	42	50	47	49	41	50	50	50	48	47
	50	50	50	50	50	50	50	49	50	50

$m' = 51$

$\hat{\hat{p}} = .705407$ with $k = 4$ iterations

$\hat{p} = .705683$ with $k = 5$ iterations

$\check{p} = .688143$

$\tilde{p} = .663025$

$\bar{p} = .704784$

Example 29) $p_o = .8$, $n = 75$, $t = 60$, $m = 25$

x_i:	60	60	60	58	56	57	54	60	58	59
	60	60	60	57	60	57	58	60	60	60
	60	53	60	60	57					

$m' = 11$

$\hat{\hat{p}} = .798565$ with $k = 4$ iterations

$\hat{p} = .798339$ with $k = 5$ iterations

$\check{p} = .780800$

$\breve{p} = .756364$

$\bar{p} = .799009$

Example 30) $p_o = .8$, $n = 75$, $t = 60$, $m = 50$

x_i:

56	58	60	60	58	60	60	56	58	60
56	60	60	60	54	60	60	60	60	60
58	60	59	55	60	60	58	60	60	60
60	60	60	60	60	60	59	60	60	57
60	60	52	56	60	56	58	60	59	60

$m' = 18$

$\hat{\hat{p}} = .807110$ with $k = 4$ iterations

$\hat{p} = .806290$ with $k = 5$ iterations

$\check{p} = .784800$

$\tilde{p} = .757778$

$\bar{p} = .808440$

Example 31) $p_o = .8$, $n = 75$, $t = 60$, $m = 75$

x_i:

60	53	60	60	52	59	56	60	60	60
60	54	60	53	60	60	58	60	60	60
58	60	58	60	60	60	56	56	57	54
60	53	60	60	59	60	60	60	60	60
60	60	55	60	60	60	57	57	60	59
60	56	60	58	53	60	55	60	60	59
60	60	60	60	52	60	60	56	55	60
57	60	60	59	48					

$m' = 30$

$\hat{\hat{p}} = .795696$ with $k = 4$ iterations

$\hat{p} = .792034$ with $k = 5$ iterations

$\check{p} = .777244$

$\tilde{p} = .743111$

$\bar{p} = .803692$

Example 32) p_o = .8, n = 75, t = 60, m = 100

x_i:	59	60	60	60	60	57	56	60	60	60
	60	57	59	59	60	60	57	55	60	60
	60	60	54	60	60	59	57	60	60	59
	60	60	56	60	60	60	60	60	59	60
	60	56	59	53	54	60	60	57	60	56
	60	57	60	57	57	58	59	60	60	60
	59	59	57	60	59	60	60	59	60	53
	60	60	58	55	60	60	60	59	60	57
	60	58	53	60	60	60	57	54	60	54
	60	56	58	59	60	57	60	60	58	60

m' = 45

$\hat{\hat{p}}$ = .799656 with k = 4 iterations

$\hat{p}$ = .800614 with k = 5 iterations

$\check{p}$ = .782000

$\tilde{p}$ = .760000

$\bar{p}$ = .797846

Example 33) p_o = .9, n = 75, t = 70, m = 25

x_i:	65	68	70	67	65	70	66	70	64	67
	65	65	70	68	69	68	68	69	68	70
	70	69	68	63	67					

m' = 19

$\hat{\hat{p}}$ = .903597 with k = 6 iterations

$\hat{p}$ = .903854 with k = 5 iterations

$\check{p}$ = .900800

$\tilde{p}$ = .890526

$\bar{p}$ = .901417

Example 34) p_o = .9, n = 75, t = 70, m = 50

x_i:	69	68	64	68	69	70	60	70	67	64
	59	70	66	70	70	66	66	66	67	70
	70	66	62	67	63	70	67	66	68	70
	64	67	66	66	70	64	70	69	69	67
	65	67	63	67	70	67	68	70	69	67

m' = 37

$\hat{\hat{p}}$ = .896911 with k = 6 iterations

$\hat{p}$ = .896218 with k = 5 iterations

$\check{p}$ = .894133

$\tilde{p}$ = .880360

$\bar{p}$ = .903611

Example 35) $p_o = .9$, $n = 75$, $t = 70$, $m = 75$

x_i:

69	68	67	67	70	69	66	70	69	68
64	66	68	64	66	68	65	67	66	65
69	64	70	64	64	65	69	66	65	70
67	66	65	70	65	68	70	70	70	67
67	65	69	68	70	66	70	68	70	66
70	70	64	60	68	67	69	68	70	66
68	65	63	69	66	70	67	66	68	67
66	66	69	70	68					

$m' = 59$

$\hat{\hat{p}} = .899237$ with $k = 6$ iterations

$\hat{p} = .899329$ with $k = 5$ iterations

$\check{p} = .896889$

$\tilde{p} = .887006$

$\bar{p} = .898219$

Example 36) $p_o = .9$, $n = 75$, $t = 70$, $m = 100$

x_i:

67	69	69	67	70	63	65	67	67	63
68	70	65	70	65	65	66	69	66	67
66	67	67	68	68	65	70	62	64	68
69	69	69	69	68	68	67	61	70	67
67	64	64	68	69	68	70	70	63	67
70	70	69	66	70	67	67	67	65	69
69	68	64	65	70	69	66	63	70	65
69	70	67	69	70	65	68	70	65	70
65	70	70	66	68	66	70	65	69	70
63	67	69	65	68	70	63	69	69	65

$m' = 79$

$\hat{\hat{p}} = .899520$ with $k = 6$ iterations

$\hat{p} = .899684$ with $k = 5$ iterations

$\check{p} = .897200$

$\tilde{p} = .887595$

$\bar{p} = .897787$

The following tables give the mean and variance for all the data sets.

p_o	n	t	m		$\hat{\hat{p}}$	$\hat{p}$	$\check{p}$	$\tilde{p}$	$\bar{p}$
.1	20	3	25	mean	.102089	.101498	.090900	.059367	.101218
				variance	.000246	.000233	.000099	.000052	.000544
			50	mean	.097648	.097483	.088450	.059853	.093716
				variance	.000077	.000087	.000043	.000053	.000301
			75	mean	.099184	.098991	.089667	.060737	.098251
				variance	.000041	.000042	.000021	.000028	.000018
			100	mean	.099312	.099581	.090000	.062628	.099242
				variance	.000065	.000065	.000032	.000017	.000422
.2	30	7	25	mean	.201875	.203042	.185467	.155444	.198696
				variance	.000307	.000327	.000125	.000119	.000379
			50	mean	.201869	.200899	.184633	.149733	.203528
				variance	.000122	.000126	.000049	.000038	.000123
			75	mean	.202349	.202964	.186022	.154936	.200772
				variance	.000084	.000095	.000037	.000047	.000083
			100	mean	.204352	.204505	.187050	.154431	.203772
				variance	.000072	.000069	.000027	.000022	.000099
.3	40	15	25	mean	.296656	.296532	.291400	.273590	.293478
				variance	.000115	.000122	.000091	.000121	.000276
			50	mean	.302502	.301887	.296050	.275761	.302010
				variance	.000108	.000105	.000078	.000083	.000186
			75	mean	.298337	.297808	.292583	.273243	.298499
				variance	.000058	.000057	.000042	.000055	.000127
			100	mean	.300817	.300342	.294800	.275436	.299878
				variance	.000041	.000036	.000027	.000018	.000107

p_o	n	t	m		$\hat{\hat{p}}$	$\hat{p}$	$\check{p}$	$\tilde{p}$	$\bar{p}$
.4	50	20	25	mean	.405644	.406932	.375240	.341198	.403810
				variance	.000192	.000192	.000039	.000141	.000299
			50	mean	.400743	.401340	.372700	.339066	.399929
				variance	.000152	.000164	.000042	.000092	.000190
			75	mean	.401011	.400614	.372560	.337059	.401950
				variance	.000074	.000088	.000022	.000060	.000076
			100	mean	.399996	.399908	.372350	.337926	.400160
				variance	.000039	.000038	.000009	.000021	.000057
.5	60	30	25	mean	.498078	.498659	.473167	.442379	.496805
				variance	.000187	.000176	.000042	.000011	.000293
			50	mean	.498285	.499024	.473450	.443433	.496837
				variance	.000142	.000142	.000031	.000052	.000201
			75	mean	.496523	.496433	.472389	.440988	.496712
				variance	.000047	.000048	.000013	.000036	.000073
			100	mean	.500422	.500831	.474583	.443784	.499787
				variance	.000059	.000062	.000015	.000015	.000061
.6	65	40	25	mean	.600844	.602552	.583538	.559164	.597359
				variance	.000231	.000275	.000094	.000156	.000255
			50	mean	.600467	.600091	.582677	.554407	.601450
				variance	.000064	.000069	.000026	.000049	.000082
			75	mean	.600323	.600326	.582933	.555604	.601892
				variance	.000028	.000029	.000011	.000032	.000074
			100	mean	.598333	.598250	.581608	.554601	.597955
				variance	.000043	.000049	.000018	.000031	.000059

p_o	n	t	m		$\hat{\hat{p}}$	$\hat{p}$	$\check{p}$	$\tilde{p}$	$\bar{p}$
.7	70	50	25	mean	.696840	.696846	.682371	.657922	.697112
				variance	.000107	.000123	.000049	.000078	.000115
			50	mean	.697639	.698190	.683414	.660231	.696036
				variance	.000063	.000063	.000025	.000039	.000099
			75	mean	.697332	.697339	.682876	.658454	.696715
				variance	.000061	.000064	.000026	.000046	.000099
			100	mean	.699684	.700283	.684793	.661357	.698406
				variance	.000033	.000039	.000015	.000027	.000035
.8	75	60	25	mean	.798669	.799539	.780667	.758339	.797405
				variance	.000165	.000191	.000049	.000069	.000167
			50	mean	.799852	.800199	.781547	.758034	.799181
				variance	.000060	.000059	.000017	.000024	.000073
			75	mean	.799266	.799811	.781502	.758253	.798231
				variance	.000017	.000019	.000005	.000034	.000039
			100	mean	.800362	.800959	.782073	.759094	.799258
				variance	.000025	.000026	.000007	.000013	.000032
.9	75	70	25	mean	.900603	.900887	.898187	.888538	.897791
				variance	.000020	.000023	.000017	.000029	.000039
			50	mean	.900724	.900499	.897867	.885783	.902112
				variance	.000017	.000017	.000013	.000020	.000057
			75	mean	.900300	.900216	.897636	.886482	.900766
				variance	.000011	.000011	.000008	.000011	.000028
			100	mean	.900091	.900061	.897507	.886720	.900095
				variance	.000009	.000009	.000007	.000007	.000023

It is clear that we can break down our discussion into 3 cases.

(i) $m' = 0$

In this case $x_i = t$ for all i, i.e. all the x_i's attain the bound.

From previous discussion, we see that $\hat{\hat{p}} = \bar{p} = 1$ which overestimates p.

$\tilde{p}$ is being undefined, $\check{p} = \frac{t}{n}$ always underestimates p. Hence the moment estimate provides a better estimate in this extreme case.

(ii) $m' = m$

In this case, $\hat{\hat{p}} = \check{p} = \tilde{p} = \frac{\bar{x}}{n}$. We also see from section 5.1 that $\frac{\bar{x}}{n}$ may well be a reasonable estimate. Finally we see that $\bar{p} = 0$ which underestimates p.

(iii) $0 < m' < m$

As we scan over all the values of the estimates in the table, we notice at once that $\check{p}$ underestimates p while that of $\tilde{p}$ even more underestimates p. The ratio estimates give a better estimator but with a larger variance among all of them. A close investigation of the maximum likelihood estimates $\hat{\hat{p}}$ and moment estimates $\hat{p}$ show that they provide the best estimates of p with respect to others. However, it seems hard to distinguish which estimate is more superior than the other.

In order to give a clear picture, we summarize the above results in the following table.

Estimator	$m' = 0$ i.e. $x_i = t$ $\forall i$	$0 < m' < m$	$m' = m$ i.e. $x_i < t$ $\forall i$
MLE $\hat{\hat{p}}$	1 overestimate	best estimate	$\sum_{i=1}^{m} x_i/nm = \bar{x}/n$
moment estimate $\hat{p}$	best estimator	best estimate	may consider $\sum_{i=1}^{m} x_i/nm = \bar{x}/n$
pooled estimate $\check{p}$	underestimate	underestimate	$\sum_{i=1}^{m} x_i/nm = \bar{x}/n$
selected estimate $\tilde{p}$	undefine	more underestimate	$\sum_{i=1}^{m} x_i/nm = \bar{x}/n$
ratio estimate $\bar{p}$	1 overestimate	better estimate	0 underestimate

CHAPTER 7. SUMMARIES AND CONCLUSIONS

In Chapter 1, we introduce some real life examples that the binomial distribution is not applicable. This is due to the fact that the occurrences of the events are bounded. As a result, this lead to the introduction of the Bounded Above Binomial Distribution which is a special case of Lam (1982) when s = 2. In the second chapter, we introduce the p.d.f. of the distribution. The probability generating function, characteristic function, moment generating function and moments are also given. In Chapter 3, we disprove the existence of sufficient statistic and minimum variance bound estimator of p. Chapter 4 deals with the discussion of maximum likelihood estimator. The consistency, efficiency and asymptotic normality of ML estimator are carefully studied. The main problem in this chapter is the uniqueness of ML estimator. Finally we give a brief discussion of the successive approximation to ML estimator using the moment estimate as the initial estimate. In the computer simulation, it is found that the ML estimator and the approximation agree excellently. For Chapter 5, we come into consideration of moment estimator, pooled estimate, selected estimate and ratio estimate. The behaviour of the various estimates in the extreme cases are also listed out. Comparison of various estimates by computer simulation is found in Chapter 6. After careful screening, it is found that the ML estimator and moment estimator are the best in compared with the others as would be expected.

REFERENCES

Hildebrand, F.B. (1956). Introduction to Numerical Analysis. McGraw Hill. N.Y./Toronto/London, pp. 447-450.

Hogg, R.V. and Tanis, E.A. (1977). Probability and Statistical Inference. Collier MacMillan International Editions.

Huzurbazar, V.S. (1948). The likelihood equation, consistency and the maxima of the likelihood function. Ann. Eugen., 14, pp. 185.

Hymans, S.H. (1967). Probability Theory with Application to Econometrics and Decision-Making. Prentice-Hall International Series in Management.

Johnson, N.L. and Kotz, S. (1969). Distribution in Statistics — Discrete Distributions. New York: John Wiley and Sons.

Kendall, M.G. and Stuart, A. (1973). The Advanced Theory of Statistics. Vol. 2. Griffin.

Lam, H.K. (1982). Bounded Multinomial Distribution. Communications in Statistics A11(16), pp. 1869-1880.

Larson, H.R. (1966). A nomograph of the cumulative binomial distribution. Industrial Quality Control, 23, pp. 270-278.

Rao, C.R. (1973). Linear Statistical Inferences and Its Applications. New York: John Wiley and Sons.